Contraception

Contraception 90 (2014) 142 – 146

Original research article

Different Pearl Indices in studies of hormonal contraceptives in the United States: impact of study population☆

Christoph Gerlinger[a,b,*], James Trussell[c,d], Uwe Mellinger[a], Martin Merz[a], Joachim Marr[a], Ralf Bannemerschult[a], Ilka Schellschmidt[a], Jan Endrikat[a,b]

[a]*Bayer HealthCare Pharmaceuticals, 13353 Berlin, Germany*
[b]*Gynecology, Obstetrics and Reproductive Medicine, University Medical School of Saarland, 66421 Homburg/Saar, Germany*
[c]*Office of Population Research, Princeton University, Princeton, NJ 08544, USA*
[d]*The Hull York Medical School, Hull HU6 7RX, UK*

Received 19 December 2013; revised 21 March 2014; accepted 23 March 2014

Abstract

Objective: To examine the impact of subject characteristics on efficacy as measured by the Pearl Index (PI) in clinical trials and to make study populations similar by matching.
Methods: Our analysis used US data from four large Phase III studies. We compared results from one fertility control patch study with pooled data from three studies with virtually identical design on oral hormonal contraceptives. First, we identified three characteristics that had the most impact on the PI. Second, we used these three variables and matched subjects from the patch study with those from the oral contraceptive (OC) studies. Finally, we calculated the PIs for matched and unmatched subjects from both the patch study and the OC studies.
Results: A total of 3706 subjects were included in our analysis. The variables 'Hispanic ethnicity', 'previous pregnancy' and 'previous use of hormonal contraceptives' had the most impact on the PI. The PIs for the matched patch cohort and the matched OC cohort were 2.97 and 2.48, respectively. Those for the unmatched patch cohort and the unmatched OC cohort were 10.17 and 0.90, respectively.
Conclusion: Subject characteristics strongly influence the PI in clinical studies of hormonal contraceptives. In particular, Hispanic ethnicity, previous pregnancies and no previous use of hormonal contraceptives result in a higher PI.
Implications: PIs from different clinical trials cannot be meaningfully compared unless subject characteristics that have most impact on the PI are similar or are made to be similar statistically as we did here by matching.

Keywords: Pearl Index; Contraception; Study population; Subject characteristics; Contraceptive failure

1. Introduction

Hormonal contraceptives are among the most popular, safe and effective methods of reversible contraception [1]. Authorities all over the world — including the US Food and Drug Administration and the European Medicines Agency — stipulate assessment of efficacy by the Pearl Index (PI). The PI is the primary endpoint of those Phase III clinical trials. Recently approved oral contraceptives (OCs) in the US featured PIs between 2 and 3 (LoSeasonique: 2.74 [2]; Lo Loestrin Fe: 2.92 [3]; Quartette: 3.19 [4]) while those developed previously showed PIs below 2 (Yasmin: 0.406 [5]; Yaz™: 1.29 [6]). Apparently, PIs appear to be increasing over time [7].

In a PI study in the US starting in 2009, a fertility control patch showed an unadjusted PI of 3.56 [upper 95% confidence interval (CI) 4.95] [8] in women aged 18–45 years with no body mass index (BMI) restrictions. We sought to explore whether this high PI could be explained by

☆ Conflict of interest and financial disclosure: C.G., U.M., M.M., J.M., R.B., I.S. and J.E. are employees of Bayer Healthcare. J.T. is a consultant for Bayer economists working on the cost effectiveness of contraception. His work was supported in part by the Eunice Kennedy Shriver National Institute of Child Health and Human Development grant for Infrastructure for Population Research at Princeton University, Grant R24HD047879.

* Corresponding author. Bayer HealthCare Pharmaceuticals, Müllerstrasse 178, 13353 Berlin, Germany. Tel.: +49 30 468 11829; fax: +49 30 468 91829.
E-mail address: christoph.gerlinger@bayer.com (C. Gerlinger).

http://dx.doi.org/10.1016/j.contraception.2014.03.018

different subject variables. Therefore, we performed a comparison with pooled data from three other hormonal contraceptive studies (Natazia/Qlaira™, Yaz™ and Flexyess™) starting between 2000 and 2007. First, we ran an exploratory analysis using all available variables related to demography, medical history and gynecological and reproductive history. Based on the evaluation of a total of 13 variables, we identified three variables that impacted PI the most. Second, we used these three variables and matched subjects from the patch study with those from the pooled OC studies by using the propensity score analysis method. Finally, we calculated the PIs for the matched and unmatched populations from both the patch study and the pooled OC studies.

2. Materials and methods

2.1. Studies included in the analysis

Four large Phase III studies sponsored by Bayer were included in this analysis. The primary efficacy outcome was contraceptive efficacy measured by the PI. Two studies were performed only in the US [8,9]; two also had centers in Europe. All lasted at least 1 year and included a total of 6602 women, of these 3706 women were treated in the US.

The most recent study investigated a fertility control patch (patch) [8] while the previous studies investigated three different oral hormonal contraceptives (OC) [6,9,10]. The estrogen component was either ethinylestradiol (EE) or estradiol valerate (E2V). The progestin component was gestodene (GSD), drospirenone (DRSP) or dienogest (DNG) in different regimens (Table 1).

2.2. Variables

The complete set of 13 variables recorded as baseline characteristics in all four studies was included in the analysis: age, race, Hispanic ethnicity (Y/N), BMI, alcohol consumption (Y/N), smoking (Y/N), previous pregnancies (Y/N), number of previous births, number of previous abortions, history of births (Y/N), history of abortions (Y/N), ever pregnant (Y/N) and previous use of hormonal contraceptives (Y/N).

2.3. Statistics

Following the intention-to-treat principle [11], all subjects from the US centers in the respective full analysis sets of the studies were retained for this analysis, provided that the demographic baseline characteristics had been recorded for all variables included in the propensity score model.

The propensity of each woman to be included in the patch study rather than in an OC study was calculated using several logistic regression models. For each subject from the patch study, one subject from OC study population was sought who had a matching propensity score. PIs were calculated separately for the matched and the unmatched subjects for the patch study and for the other OC studies.

All analyses were exploratory ex post analyses. Descriptive statistics of the data set were performed with version 9.2 of SAS software [12]. The propensity scores were calculated using version 2.13.1 of R software [13], the R-function "pscore" of the R-Package "non-random" [14], and the propensity score matching was performed using the R-function "ps.match" of the R-Package "non-random" using a caliper size of 0.01. The CIs for the PIs were calculated using the Poisson model according to Gerlinger et al. [15] and were calculated based on all pregnancies regardless of possible user intake errors.

3. Results

A total of 6602 subjects were included in all four studies worldwide (Table 1). A subset of 3706 subjects were from the US and used for this analysis. Of these, 1453 subjects were recruited from the patch study and 2253 subjects were recruited from the pool of subjects from the OC studies (Table 2).

We identified three out of 13 variables as relevant matching criteria: (1) Hispanic ethnicity (Y/N), (2) previous pregnancies (Y/N) and (3) previous use of hormonal contraceptives (Y/N). By applying these three matching criteria, 1386 subjects could be matched for one of the eight possible combinations. Sixty-seven subjects from the patch study and 867 from the OC studies could not be matched. In the matched population, 254 subjects (18.3%) were of Hispanic ethnicity, 683 subjects (49.3%) were previously

Table 1
Studies included in the analysis

Author	Contraceptive preparation	Country	Women started, *n*	Reference
Merz 2012	Patch: 0.55 mg EE/2.1 mg GSD/patch, 3×7 days, 7 days off, APLEEK™	USA	1454	[8]
Bachmann 2004	Pill: 20 μg EE/3 mg DRSP; 24 days, 4 days off, YAZ™	Austria, Argentina, Brazil, Poland, USA	1018	[6]
Nelson 2013	Pill: E2V/DNG; 5-phasic regimen, no break, QLAIRA/NATZIA™	USA, Canada, Europe	2266	[10]
Jensen 2012	Pill: 20 μg EE/3 mg DRSP; flexible regimen for 24–120 days, 4 days off in case of bleeding, Flexyess™	USA	1864	[9]

Table 2
Demographics by study type and matching status

	Patch study before matching	OC studies before matching	Patch not matched	Patch matched	OC matched	OC not matched
Number of women	1453	2253	67	1386	1386	867
Age (mean±S.D.)	27.9±6.4	25.6±5.1	29.5±7.2	27.9±6.4	25.5±4.9	25.6±5.4
BMI (mean±S.D.)	26.8±6.2	24.0±3.9	28.0±6.6	26.8±6.1	24.0±3.9	24.2±3.9
Race						
Caucasian	852 (58.6%)	1654 (73.4%)	13 (19.4%)	839 (60.5%)	918 (66.2%)	736 (84.9%)
Black	225 (15.5%)	187 (8.3%)	6 (9.0%)	219 (15.8%)	145 (10.5%)	42 (4.8%)
Hispanic	301 (20.7%)	314 (13.9%)	47 (70.1%)	**254 (18.3%)**	**254 (18.3%)**	60 (6.9%)
Asian	38 (2.6%)	55 (2.4%)	**0**	38 (2.7%)	37 (2.7%)	18 (2.1%)
Other	37 (2.5%)	43 (1.9%)	1 (1.5%)	36 (2.6%)	32 (2.3%)	11 (1.3%)
Ever pregnant	743 (51.1%)	885 (39.3%)	60 (89.6%)	**683 (49.3%)**	**683 (49.3%)**	202 (23.3%)
Hormonal contraceptives before study	600 (41.3%)	1286 (57.1%)	0	**600 (43.3%)**	**600 (43.3%)**	686 (79.1%)

Boldface, criteria for matching.

pregnant and 600 subjects (43.3%) used hormonal contraceptives prior to study start for each of the patch and the OC users group. In the unmatched patch cohort, 70.1% of subjects were of Hispanic ethnicity and 89.6% were previously pregnant compared to 6.9% and 23.3% in the unmatched OC group. In the unmatched OC cohort, 686 subjects (79.1%) used hormonal contraceptives before the study in contrast to 0% in the unmatched patch cohort (Table 2). Thus, Hispanic women and subjects who have been previously pregnant were more likely to be included in the patch study while subjects who previously used hormonal contraceptives were more likely included in an OC study.

The PIs of the studies before matching were 3.26 (CI 2.23–4.6) for the patch study and 1.85 (CI 1.62–2.62) for the OC studies. The PIs for the matched patch cohort and the matched OC cohort were 2.97 (CI 1.97–4.29) and 2.48 (CI 1.60–3.66), respectively. Those for the unmatched patch cohort and the unmatched OC cohort were 10.17 (CI 2.77–26.05) and 0.90 (CI 0.33–1.96), respectively (Table 3).

4. Discussion

PIs appear to be increasing over time [7] as also reflected in the high PI for the fertility control patch investigated here. While Trussell and Portman proposed two main factors for increasing PIs, first, the more frequent pregnancy testing with more sensitive tests, and, second, the decrease in adherence due to changes in study populations over time [7], we focused specifically on the latter point. We felt it was worthwhile to identify factors that might explain the rise in PIs.

We identified three relevant subject variables as important factors impacting the PI: (1) Hispanic ethnicity, (2) previous pregnancies and (3) previous use of hormonal contraceptives. After matching the fertility control patch population with the pooled OC population for these three subject variables, the PIs for the two matched cohorts were equal. We would suggest that this is strong evidence that the PI of 3.56 in the patch study is higher because of the characteristics of the specific study population; specifically, a high percentage of Hispanic subjects, women that gave birth or had an abortion before and women who rarely used effective hormonal contraception before resulted in a higher contraceptive failure rate.

There is some evidence in the literature about the impact of the study population on the PI. For example, a recent study comparing an experimental patch to a standard pill containing 20 μg EE and 100 μg levonorgestrel also showed PIs of approximately 5 and 4 [16] in a comparable population as that of the patch study presented here.

The overall PI summarizes pregnancies due to method and subject failures. While method failures occur during ‘perfect use’, subject failures occur during ‘typical use’. ‘Typical use’ reflects how effective methods are for the

Table 3
Pearl Indices

	Patch study before matching	OC studies before matching	Patch not matched	Patch matched	OC matched	OC not matched
Relevant exposure during complete study (women-years)	982.6	1674.0	39.3	943.3	1009.4	664.6
Pregnancies[a], *n* (%)	32 (2.2)	31 (1.4)	4 (6.0)	28 (2.0)	25 (1.8)	6 (0.7)
Pearl Index	3.26	1.85	10.17	2.97	2.48	0.90
95% CI for PI	2.23–4.60	1.26–2.62	2.77–26.05	1.97–4.29	1.60–3.66	0.33–1.96

[a] During treatment and until 7 days after last patch wear or pill intake.

average person who does not always use methods correctly or consistently [17]. Thus, subject failures are caused by non-compliance. Westhoff et al. investigated predictors of non-compliance including race, previous pregnancies and current OC use. In the group of non-compliant subjects, significantly more Hispanics and blacks as well as previously pregnant women were found. Current OC use was numerically but not significantly higher in the group of compliant subjects [18]. Our findings on pregnancies in these populations are well in line. In Westhoff's logistic regression analysis, lack of education, residential poverty, Hispanic ethnicity and obesity continued to be associated with non-compliance. Westhoff et al. suggest that race or ethnicity (black/Hispanic) does not directly impact contraceptive failure rates but more likely is a proxy for socioeconomic parameters such as low income, low educational level, lack of insurance coverage, etc. [18].

Also, Borrero et al. described the impact of race/ethnicity on differences in compliance. In a cohort of 6946 female veterans, Hispanics were significantly more likely than whites to experience gaps between refills of contraceptives. In addition, Hispanic and black women had fewer months of contraceptive coverage than white women [19]. In a study on compliance comparing a contraceptive patch versus an OC, Kaunitz et al. found higher rates of perfect compliance in patch users independent of age, race/ethnicity and education. In the group of OC users, all these factors influenced perfect compliance with lower values in the group of Hispanics and in women with lower education [20]. In addition to Hispanic ethnicity, we found that a history of previous pregnancies increases the likelihood of contraceptive failures. This finding is in line with that in Dinger et al., although they also found a positive correlation between contraceptive failures and lower education, higher BMI and young age (20–24 years) [21].

Interestingly, Kost et al. reported estimates of contraceptive failure from the 2002 National Survey of Family Growth. They concluded that the risk of failure is strongly affected by socioeconomic characteristics of the users, e.g., probability of failure higher for Hispanic and black women than for whites, more than twice as high for women with one or more births as for those with none, more than twice as high for those with lower incomes than for those with higher incomes and higher for cohabiting women than for married women [22]. A similar pattern is seen for the incidence of unintended pregnancies. Finer et al. found that unintended pregnancies are highest in black or Hispanic women and in women with low education (no high school graduation), with low income (<100% of poverty) and with ≥2 births [23]. However, it is unclear whether unintended pregnancy is due to no use of any kind of contraception or to imperfect use. Finally, geography also has some influence on the PI. When one hormonal contraceptive is investigated in American and European population, the PIs derived from US are usually higher than those from Europe (for example, Flexyess™: Europe, 0.64 [24] versus USA, 1.65 [9]). This phenomenon needs further investigation.

One limitation of our analysis is that, in all four studies, we did not specifically ask whether women wanted a child in the future. Whether they were spacers or limiters could well have affected compliance and eventually rates of contraceptive failure.

Finally, the analysis presented here was conducted ex post. A confirmation of these findings in a prospective setting would be needed.

5. Conclusion

Subject variables strongly influence the PI in clinical studies of hormonal contraceptives. In particular, Hispanic ethnicity, previous pregnancies and no previous use of hormonal contraceptives result in a higher PI.

Acknowledgment

This work was supported in part by the Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant for Infrastructure for Population Research at Princeton University, Grant R24HD047879 (J.T.).

References

[1] Burkman RT. Clinical pearls: factors affecting reported contraceptive efficacy rates in clinical studies. Int J Fertil Womens Med 2002;47:153–61.
[2] Orleans RJ, Center for Drug Evaluation and Research, Application number 22-262, LoSeasonique™, October 23, 2008.
[3] Center for Drug Evaluation and Research, Application number 022501Orig1s000. Lo Loestrin Fe, October 4, 2010.
[4] [No authors listed] Quartette: an ascending-dose, extended-cycle oral contraceptive. Med Lett Drugs Ther 2013;55(1420):54–5.
[5] Parsey KS, Pong A. An open-label, multicenter study to evaluate Yasmin, a low-dose combination oral contraceptive containing drospirenone, a new progestogen. Contraception 2000;61:105–11.
[6] Bachmann G, Sulak PJ, Sampson-Landers C, Benda N, Marr J. Efficacy and safety of a low-dose 24-day combined oral contraceptive containing 20 micrograms ethinylestradiol and 3 mg drospirenone. Contraception 2004;70:191–8.
[7] Trussell J, Portman D. The creeping pearl: why has the rate of contraceptive failure increased in clinical trials of combined hormonal contraceptive pills? Contraception 2013;88(5):604–10.
[8] Data on file: Bayer Clinical Study Report No. A50132, April 2012. Multi-center, open-label, uncontrolled study to investigate the efficacy and safety of the transdermal contraceptive patch containing. 0.5 mg ethinyl estradiol and 2.1 mg gestodene (Material No. 80876395) in a 21-day regimen for 21 cycles in 1650 healthy female subjects. Bay no. 86 5016/Impact no. 91555.
[9] Jensen JT, Garie SG, Trummer D, Elliesen J. Bleeding profile of a flexible extended regimen of ethinylestradiol/drospirenone in US women: an open-label, three-arm, active-controlled, multicenter study. Contraception 2012;86:110–8.
[10] Nelson A, Parke S, Makalova D, Serrani M, Palacios S, Mellinger U. Efficacy and bleeding profile of a combined oral contraceptive containing oestradiol valerate/dienogest: a pooled analysis of three studies conducted in North America and Europe. Eur J Contracept Reprod Health Care 2013;18(4):264–73.
[11] Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319(7211):670–4.

[12] www.sas.com [accessed 2013–09–30].
[13] http://cran.r-project.org/ [accessed 2013–09–30].
[14] http://cran.r-project.org/src/contrib/Archive/nonrandom/ [accessed 2013–09–30].
[15] Gerlinger C, Endrikat J, van der Meulen EA, Dieben TOM, Düsterberg B. Recommendation for confidence interval and sample size calculation for the Pearl Index. Eur J Contracept Reprod Health Care 2003;8(2):87–92.
[16] Kaunitz AM, Mishell Jr DR, Foegh M. Comparative phase 3 study of AG200-15, a low-dose estrogen and levonorgestrel contraceptive patch. Presented at the 60th Annual Clinical Meeting of the American College of Obstetricians and Gynecologists, May 5–9, 2012, San Diego, CA; 2012. [Available at: http://www.agiletherapeutics.com/pdf/posters/ACOG%20CL12%20Poster%20FINAL.pdf].
[17] Trussell J. Contraceptive failure in the United States. Contraception 2011;83:397–404.
[18] Westhoff CL, Torgal AT, Mayeda ER, Shimoni N, Stanczyk FZ, Pike MC. Predictors of noncompliance in an oral contraceptive clinical trial. Contraception 2012;85:465–9.
[19] Borrero S, Zhao X, Mor MK, Schwarz EB, Good CB, Gellad WF. Adherence to hormonal contraception among women veterans: differences by race/ethnicity and contraceptive supply. Am J Obstet Gynecol 2013;209:103.e1–03.e11.
[20] Kaunitz A, Foegh D, Mishell M. Increased compliance independent of patient demographics with a low-dose combination contraceptive patch (AG200-15) compared with a low-dose combination oral contraceptive in a phase 3 clinical trial. Contraception 2012;86:314.
[21] Dinger J, Minh TD, Buttmann N, Bardenheuer K. Effectiveness of oral contraceptive pills in a large U.S. cohort comparing progestogen and regimen. Obstet Gynecol 2011;117:33–40.
[22] Kost K, Singh S, Vaughan B, Trussell J, Bankole A. Estimates of contraceptive failure from the 2002 National Survey of Family Growth. Contraception 2008;77:10–21.
[23] Finer LB, Zolna MR. Unintended pregnancy in the United States: incidence and disparities, 2006. Contraception 2011;84:478–85.
[24] Klipping C, Duijkers I, Fortier MP, Marr J, Trummer D, Elliesen J. Contraceptive efficacy and tolerability of ethinylestradiol 20 μg/drospirenone 3 mg in a flexible extended regimen: an open-label, multicentre, randomised, controlled study. J Fam Plann Reprod Health Care 2012;38:73–83.